

Analisis Penerapan *Audio Ducking* Dalam Paska Produksi Multimedia dengan Metode *Basic* dan *Multiband Sidechain Compression*

Maulana Brama Shandy¹, Ema Utami², Dhani Ariatmanto³

Teknik Informatika, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia
Corresponding Author: maulana.shandy@students.amikom.ac.id

Abstract

This study addresses the issue of inaudible speech in multimedia content due to the lack of post-production techniques, particularly in the ducking process, which remains a persistent problem in the audio engineering industry. This study presents two methods of sidechain compression: basic sidechain compression and multiband compression, aimed at improving the quality of speech in multimedia content. Using a comparative analysis of the two compression methods, this paper recommends the most suitable method for each musical genre that will be used in multimedia content. This research offers valuable insights for media industry professionals and audio processing researchers, as it provides a better understanding of the role of sidechain compression in enhancing the quality of multimedia content. The results of the study demonstrate the effectiveness of the compression methods in improving the quality of dialog in multimedia content.
Keywords: *Audio Processing; Sidechain Compression; Speech intelligibility; Audio Ducking; Multimedia*

Abstrak

Studi ini membahas masalah ketidakjelasan dialog pada konten multimedia karena kurangnya teknik pasca-produksi, terutama pada proses *ducking*, yang masih menjadi masalah yang persisten dalam industri rekayasa audio. Studi ini mempresentasikan dua metode *sidechain compression: basic sidechain compression dan multiband compression*, yang bertujuan untuk meningkatkan kualitas pidato dalam konten multimedia. Dengan melakukan analisis perbandingan dari kedua metode kompresi, tulisan ini merekomendasikan metode yang paling cocok untuk setiap genre musik yang akan digunakan dalam konten multimedia. Penelitian ini memberikan wawasan berharga bagi profesional industri media dan peneliti pemrosesan audio, karena memberikan pemahaman yang lebih baik tentang peran *sidechain compression* dalam meningkatkan kualitas konten multimedia. Hasil penelitian menunjukkan efektivitas metode kompresi dalam meningkatkan kualitas dialog/narasi dalam konten multimedia.

Kata kunci: *Pemrosesan Suara; Sidechain Compression; Kejelasan Dialog; Audio Ducking; Multimedia*

1. Pendahuluan

Banyaknya tayangan multimedia yang tersedia pada media streaming membuat masyarakat dapat menikmati kreativitas para *content creator* berupa content hiburan, tutorial, *daily vlog* yang disajikan secara apik. Namun, terkadang penonton tidak dapat mendengar atau menangkap informasi dalam dialog atau narasi yang disampaikan karena terganggu dengan suara musik atau efek suara yang ada dalam tayangan tersebut terlebih jika dalam bahasa asing [1]. Gangguan tersebut sebagian besar diakibatkan oleh musik yang digunakan, jenis musik mempengaruhi seberapa besar fokus penonton dalam memperoleh informasi pada tayangan tersebut[2].

Kekurangan penonton dalam mendengar atau menangkap informasi dalam dialog atau narasi yang disampaikan dapat diasumsikan bahwa dalam tahapan proses produksi tayangan (*Pre-production, production, post-production*) kurang maksimal dikarenakan pengaruh *Loudness Difference (LD)* pada proses *Post-Production*. *Loudness Difference* adalah perbedaan tingkat kenyaringan antara suara utama dengan suara latar belakang [3]. LD dapat diaplikasikan pada proses *Sound Mixing* dalam tahapan proses *Post-Production*.

Sound Mixing adalah pencampuran suara dari beberapa suara yang digabungkan menjadi satu channel atau lebih, dalam hal ini adalah dialog, efek suara dan musik. Produksi perekaman modern dibagi menjadi tiga fasa yaitu perekaman (*Recording*), perekaman ulang (*Overdub*), dan menggabungkan semua sumber audio menjadi satu atau beberapa channel (*Mixdown*) [4]. Dalam proses pencampuran suara inilah teknik *audio ducking* digunakan dengan menggunakan *sidechain compression*.

Compression adalah proses penjagaan sinyal suara dalam rentang tingkat kenyaringan (*Volume range*). Menentukan tingkat kenyaringan suara yang diinginkan tidak terlalu rendah maupun terlalu tinggi [1]. Dalam kompresi perlu disesuaikan *attack time* (kapan kompresi berfungsi), *release time* (kapan kompresi berhenti), dan *threshold* (range kompresi tersebut).

Penerapan kompresi tersebut dilakukan menggunakan plugin yang berjalan pada perangkat lunak DAW (*Digital Audio Workstation*), DAW adalah perangkat lunak yang digunakan untuk perekaman suara, penyuntingan suara dan membuat suara. DAW menggunakan komputer sebagai host dan terintegrasi dengan *soundcard* sebagai converter audio dari analog menjadi digital maupun sebaliknya [5]. Dalam DAW kebanyakan sudah terdapat plugin untuk menerapkan *Sidechain Compression* yaitu menggunakan plugin *Compressor*.

Penggunaan teknik kompresi mungkin akan membuat sebagian besar *content creator* mengalami kesulitan karena minimnya pengetahuan tentang dasar *digital audio* dan lebih banyak fokus pada visual, maka dibutuhkan suatu rekomendasi dalam proses *audio ducking* pada fasa *post-production*, rekomendasi ini terkait dengan jenis musik, jenis kompresi dan nilai rasio kompresi yang akan digunakan dalam pembuatan konten multimedia.

Metode *Audio Ducking* telah digunakan untuk penelitian dengan topik sejenis seperti pada penelitian [3], [6]–[12]. Pada penelitian tersebut metode *Audio Ducking* digunakan untuk meningkatkan presentase informasi yang dapat diterima oleh penonton dari sebuah tayangan. Tujuan dari penelitian ini adalah mendapatkan bukti pengaruh jenis musik, jenis kompresi dan nilai rasio terhadap *Loudness Difference* (LD) sebagai acuan tingkat kejelasan suara serta rekomendasi nilai rasio berdasarkan jenis musik yang digunakan dalam proses pencampuran suara pada paska produksi multimedia.

2. Tinjauan Pustaka

Penelitian tentang tingkat penangkapan suara dan metode *ducking* telah dilakukan dengan beragam cara. Mengukur tingkat kejelasan bicara pada pendengar yang menggunakan bahasa non lokal (*non-native*) dan mengemukakan bahwa membutuhkan 1-7 dB terhadap *speech to noise ratio* untuk mencapai 50% keberhasilan penangkapan informasi oleh pendengar non bahasa lokal. Penelitian tersebut menonjolkan aspek dialog lebih penting daripada elemen lainnya seperti efek suara dan musik latar, namun pada era saat ini efek suara dan music latar menjadi faktor penunjang penyampaian informasi[6].

Menerapkan *Adjustment / Satisfaction Test*, dengan tes tutup mata dalam penyuntingan kesetaraan tingkat kenyaringan dari dialog dan musik latar, dimana penelitian ini menghasilkan tingkat perbandingan yang ideal atau disebut *Dialogue Enhancement*. Penelitian mereka menggunakan peningkatan tingkat kenyaringan pada ruang yang sama, namun pada era teknologi *mobile* sekarang ini menjadikan para pengelola *media streaming* melakukan pembatasan tingkat kenyaringan pada tayangan melalui *platform* mereka sehingga dapat menjadi acuan baru dalam menentukan tingkat kenyaringan hasil produksi tayangan[12].

Penelitian tentang SNR (*Signal to Noise Ratio*). Tujuan penelitian mereka adalah menemukan ukuran SNR (*Signal to Noise Ratio*) dalam sebuah dialog. Mereka mengemukakan bahwa ada perbedaan temuan mereka dengan standar implementasi *noise gating* 48% dalam *real time processing* dalam *speech separation*[7]. Hasil penelitian mereka dapat menjadi acuan karena menggunakan objek observasi orang yang telah mengalami penurunan pendengaran, hal ini dapat menambah referensi tingkat kenyaringan yang ideal untuk semua kalangan.

Menemukan ukuran SBR (*Signal to Background Ratio*) yang ideal untuk memenuhi kebutuhan pendengar menangkap kejelasan dialog tanpa mengorbankan musik latar dan efek suara sehingga pendengar tetap dapat menikmati sajian suara yang nyaman dalam *broadcasting*, *voice over ip* dan alat bantu dengar bagi orang yang memiliki pendegaran yang kurang[11]. Penelitian mereka dapat menjadi acuan karena memiliki kesamaan dalam tujuan penyampaian informasi tanpa mengorbankan aspek suara efek dan musik latar.

Menggunakan metode *manual volume automation* dalam *digital audio workstation* untuk melakukan proses *ducking* sehingga dialog yang dihasilkan lebih jelas untuk keperluan acara TV

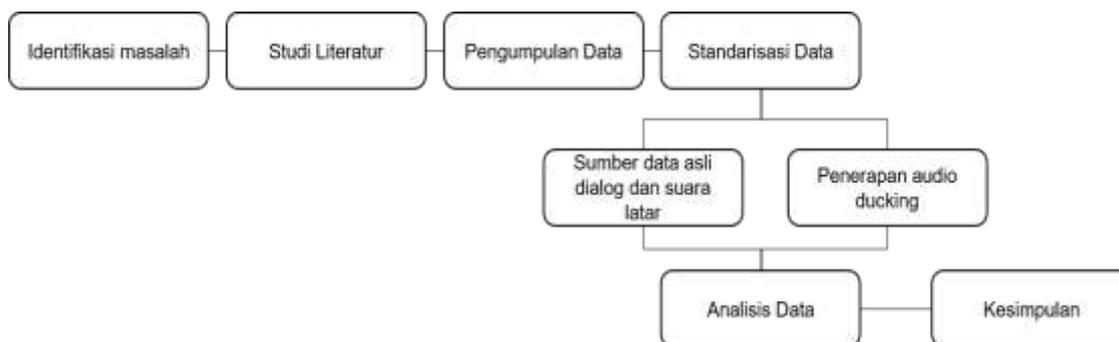
berdasarkan LD (*Loudness Different*) dengan pengukuran LU (*Loudness Units*)[8]. Kemudian melanjutkan penelitian sebelumnya mengemukakan bahwa *ducking* tidak berdasar pada buku acuan atau rekomendasi penayang akan tetapi berdasar pada dialog yang dapat dipahami dan jelas. Hasil penelitian mereka merepresentasikan bahwa sumber suara yang baik memiliki pengaruh yang besar agar dapat dipahami penonton, sumber suara tersebut meliputi pengisi suara dialog dan naskah yang akan disampaikan[9]. Berdasar penelitian mereka dapat dimodifikasi menggunakan metode yang lebih baru menggunakan *sidechain compression* untuk melakukan *ducking*.

Beberapa penelitian tersebut memiliki persamaan dengan penelitian yang akan dilakukan yaitu tema yang di teliti, memiliki persamaan yang berhubungan dengan tingkat kejelasan dialog atau narasi. Sedangkan perbedaannya yaitu mengenai metode dan tujuan. Penelitian yang akan dilakukan dikhususkan dalam produksi tayangan multimedia dimana fokus dari penelitian adalah pada analisis penerapan *audio ducking* berupa pengaruh jenis musik, metode, dan nilai rasio terhadap *Loudness Difference* (LD) yang menjadi acuan para peneliti sebelumnya untuk menentukan tingkat kejelasan suara, dengan menggunakan standar ITU-R BS.1770-4[13].

3. Metodologi

3.1 Tahapan Penelitian

Pada gambar 1 tahapan penelitian dimulai dari mencari permasalahan yang sedang terjadi dalam proses produksi audio. Kemudian melakukan studi literatur berdasarkan penelitian terdahulu dan menentukan metode yang digunakan yaitu deskriptif[14] kuantitatif[15].



Gambar 1 Tahapan Penelitian

3.2 Pengumpulan Data

Metode pengumpulan data menggunakan metode eksperimen untuk menemukan hubungan sebab akibat antara satu atau lebih variabel[16]. Sumber data yang digunakan berupa dialog dan musik dimana data dialog diperoleh dari salah satu proyek iklan MSV Studio tentang ajakan menonton film Indonesia. Sedangkan musik sebagai latar belakang diperoleh dari pustaka Youtube secara acak sesuai jenis musik yang telah di tentukan yaitu *ambient, jazz, pop, RnB, rock, cinematic*. Pemilihan berdasar durasi +-1 menit.

3.3 Standarisasi Data

Data yang diperoleh kemudian di modifikasi sehingga memenuhi standar untuk penelitian ini. Seluruh data musik disesuaikan dengan data dialog yang memiliki resolusi 24/48kHz dan tipe data .Wav menggunakan DAW Presonus Studio One, kemudian melakukan koreksi pada *Integrated Loudness* (IL) menggunakan Nugen LM Correct dengan preset ITU-R BS.1770-4[13] dengan nilai (LU) -24[17]. Mengesampingkan nilai *True Peak* (dBTP) karena dalam penelitian berfokus pada (LUFS).

Setelah data sesuai dengan standarisasi penelitian maka selanjutnya adalah memasukkan data tersebut kedalam DAW Presonus Studio One dan memotong durasi menjadi 5 detik sebagai sampel data yang akan diukur. Untuk detail sumber data dapat dilihat pada Tabel 1.

Tabel 1 Sumber data

Data Audio	Asli			Standarisasi		
	IL (LUFS)	Peak (dBTP)	jenis data	IL (LUFS)	Peak (dBTP)	jenis data
Dialog	-30.2	-12.4	.Wav	-24	-6.3	.Wav
Ambient	-10.20	0.00	.mp3	-24	-13.8	.Wav
Jazz	-13.00	-2.00	.mp3	-24	-13	.Wav
Pop	-12.20	-1.00	.mp3	-24	-12.8	.Wav
RnB	-6.90	-0.10	.mp3	-24	-17.2	.Wav
Rock	-10.30	0.10	.mp3	-24	-13.6	.Wav
Sinematik	-12.90	-0.10	.mp3	-24	-11.2	.Wav

3.4 Penerapan *Audio Ducking*

Membuat kanal baru (BUS) sebagai *background* untuk meletakkan perangkat lunak yang digunakan dan menerima sinyal dari kanal musik yang telah di tempatkan sebelumnya. *Sidechain Compression* akan berfungsi saat sinyal dari kanal dialog diterima oleh kanal BUS tersebut.

Metode *audio ducking* yang digunakan adalah metode *Basic Sidechain Compression* menggunakan perangkat lunak *compressor Fabfilter Pro C2* sebagai alat *ducking*, Teknik ini mengurangi tingkat kenyaringan secara menyeluruh pada musik yang digunakan dan metode *Multiband Sidechain Compression* menggunakan perangkat lunak *compressor Presonus Multiband Dynamic* dimana tingkat kenyaringan yang dikurangi berdasarkan frekuensi pada musik yang digunakan. Pada teknik ini dicari frekuensi dialog yang dominan dengan *bantuan EQ Analyzer* didapatkan bahwa data dialog berada pada 175 Hz – 10 kHz. Frekuensi tersebut menjadi acuan saat melakukan teknik ini.

Parameter yang digunakan pada kedua teknik *sidechain* seperti pada Tabel 2

Tabel 2 Parameter Proses *Ducking*

Parameter	Nilai
Threshold	-40 dB
Attack	50 ms
Release	500 ms
Hold	0 ms

Pengujian dilakukan satu persatu dari jenis musik dan nilai rasio untuk mendapatkan data musik yang telah terkompres untuk nantinya dijadikan bahan analisis. Hasil dari pengujian *sidechain compression* dengan menggunakan kedua teknik tersebut kemudian dihitung *Loudness Difference* (LD) menggunakan MATLAB dengan dasar perhitungan *Integrated Loudness* yang telah tersedia di MATLAB[18].

4. Hasil dan Pembahasan

4.1 Hasil Pengukuran *Loudness Difference*

Metode pengujian dengan menerapkan dua teknik *sidechain compression* dan menghitung *Loudness Difference* (LD) menghasilkan data seperti pada Tabel 3

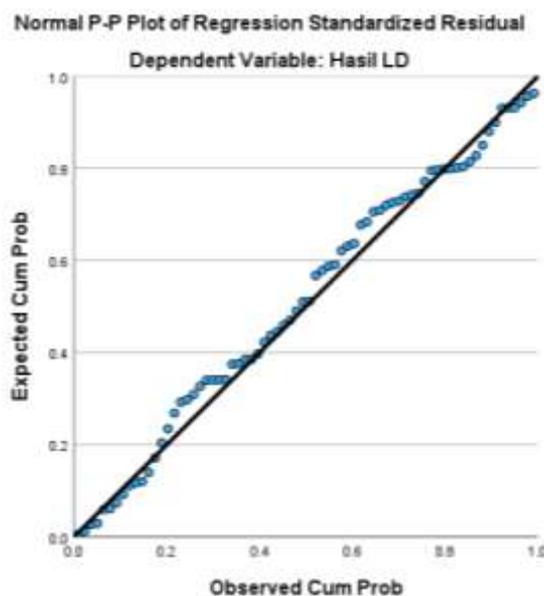
Tabel 3 Hasil pengukuran LD

Metode	Nilai rasio	Jenis Musik					
		Ambient	Jazz	Pop	RnB	Rock	Sinematik
Basic Sidechain Compression	Rasio 1:1	4.94	0.40	3.11	0.37	2.16	2.08
	Rasio 2:1	9.72	5.86	8.47	4.96	6.58	7.02
	Rasio 4:1	11.67	8.32	10.91	6.86	8.36	9.17
	Rasio 6:1	12.24	9.07	11.71	7.42	8.87	9.82
	Rasio 8:1	12.50	9.44	12.10	7.69	9.12	10.13
	Rasio 10:1	12.65	9.65	12.33	7.84	9.26	10.32
Multiband Dynamic Compression	Rasio 1:1	4.94	0.40	3.03	0.27	2.16	2.08
	Rasio 2:1	12.24	8.83	4.19	1.11	8.00	9.75
	Rasio 4:1	14.41	11.50	4.35	1.28	9.28	12.15
	Rasio 6:1	14.93	12.15	4.39	1.37	9.55	12.74
	Rasio 8:1	15.16	12.43	4.40	1.38	9.66	13.01
	Rasio 10:1	15.29	12.59	4.41	1.39	9.73	13.16

Hasil pengukuran LD pada Tabel 3 diatas menjadi acuan sebagai data yang dianalisa menggunakan ANOVA yang telah memenuhi kriteria analisis tersebut sehingga di dapatkan data yang valid sebagai alat penduga.

4.2 Uji Kelayakan Model

Untuk menentukan ketepatan model dilakukan pengujian normalitas sehingga data tersebut valid sebagai alat penduga pada analisis regresi. Pengujian menggunakan uji *probability plot* sebagai dasar pengambilan keputusan. Hasil dari pengujian normalitas tersebut seperti yang ditunjukkan pada gambar

Gambar 2 Grafik uji normalitas *probability plot*

Dari hasil tersebut dapat diketahui bahwa data tersebut valid dan dapat digunakan sebagai data penduga karena model regresi berdistribusi normal. Kemudian dilakukan uji F dan uji T untuk mengetahui pengaruh jenis musik, jenis kompresi dan nilai rasio kompresi terhadap *Loudness Difference (LD)*. Hasil uji tersebut dapat dilihat pada Tabel 4 berikut

Tabel 4 Hasil uji F dan T pada analisis regresi

Model	Uji F		Uji t	
	F	Sig.	t	Sig.
Pengaruh Jenis Musik terhadap LD			-2.048	0.044
Pengaruh Nilai Rasio terhadap LD	11.298	<.001 ^b	5.426	<.001
Pengaruh Jenis Kompresi terhadap LD			-0.51	0.612

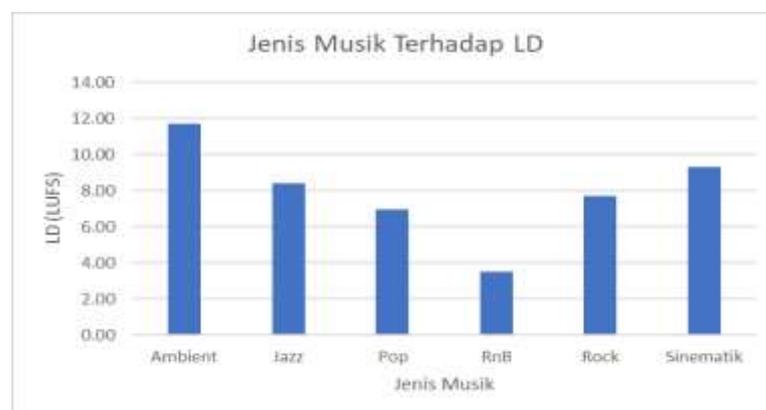
Dalam pengujian ini variabel independent yang digunakan adalah Jenis Musik (X1), Nilai rasio (X2) dan Jenis Kompresi(X3) sedangkan variabel dependen yang digunakan adalah hasil pengukuran *Loudness Difference* (Y).

Untuk mengetahui pengaruh jenis musik, jenis kompresi dan nilai rasio terhadap hasil LD digunakan analisis varian (ANOVA) karena data yang diteliti memiliki 3 variabel independent. Dari hasil tabel dapat diketahui bahwa nilai signifikansi (Sig.) dalam uji F ini sebesar $.001^b < 0,05$. Berdasarkan hal tersebut maka dapat disimpulkan bahwa Jenis Musik, Jenis Kompresi dan Nilai rasio berpengaruh secara simultan terhadap hasil LD[19].

Sedangkan untuk mengetahui pengaruh secara masing-masing (parsial) digunakan uji (t), jika nilai Sig. $< 0,05$ maka variabel independent secara parsial berpengaruh terhadap variabel dependent[19]. Pada Tabel 4 dapat diketahui bahwa pengaruh jenis musik dan nilai rasio terhadap hasil Loudness Difference (LD) dengan masing-masing nilai signifikansi (Sig.) 0.044 dan $<.001$ menunjukkan bahwa jenis musik dan nilai rasio berpengaruh secara parsial terhadap hasil LD dan signifikan. Sedangkan pengaruh jenis kompresi terhadap hasil Loudness Difference (LD) memiliki nilai signifikansi (Sig.) 0.612 yang menunjukkan bahwa jenis kompresi tidak berpengaruh secara parsial terhadap hasil LD.

4.3 Pengaruh Jenis Musik Terhadap Hasil *Loudness Difference*

Jenis musik menunjukkan nilai signifikansi (Sig.) $0,044 < 0,05$ yang berarti jenis musik berpengaruh terhadap hasil LD walaupun setiap sampel musik telah di setarakan *Loudness Unit* nya menjadi -24 LUFS berdasarkan tabel dalam standarisasi data yang digunakan untuk eksperimen. Pengaruh jenis musik terhadap LD dapat dilihat pada gambar 3

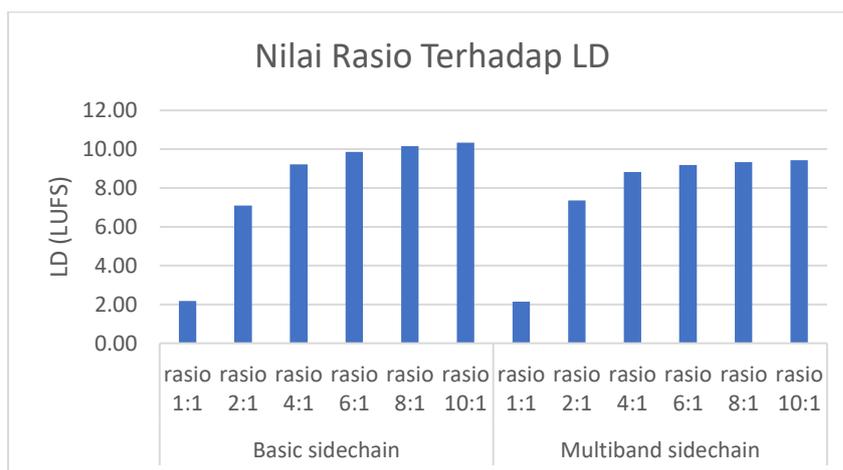


Gambar 3 Rata-rata hasil kompresi tiap jenis musik

Dari gambar 4 tersebut diketahui bahwa rata-rata hasil LD memiliki nilai berbeda-beda dari tiap jenis musik. Hal ini dapat menjadi pertimbangan konten kreator dalam memilih jenis musik yang akan digunakan dalam tayangan multimedia.

4.4 Pengaruh Nilai Rasio Terhadap Hasil *Loudness Difference*

Nilai rasio menunjukkan nilai signifikansi $< .001 < 0,05$ yang berarti nilai rasio memiliki pengaruh terhadap besaran nilai LD. Perbedaan nilai rasio (1:1, 2:1, 4:1, 6:1, 8:1, 10:1) memberikan penurunan *loudness* dari musik/latar belakang yang menjadi target *ducking* sehingga suara dialog/latar depan menjadi terdengar lebih jelas. Pengaruh nilai rasio dapat dilihat pada gambar 4



Gambar 4 Pengaruh nilai rasio terhadap LD

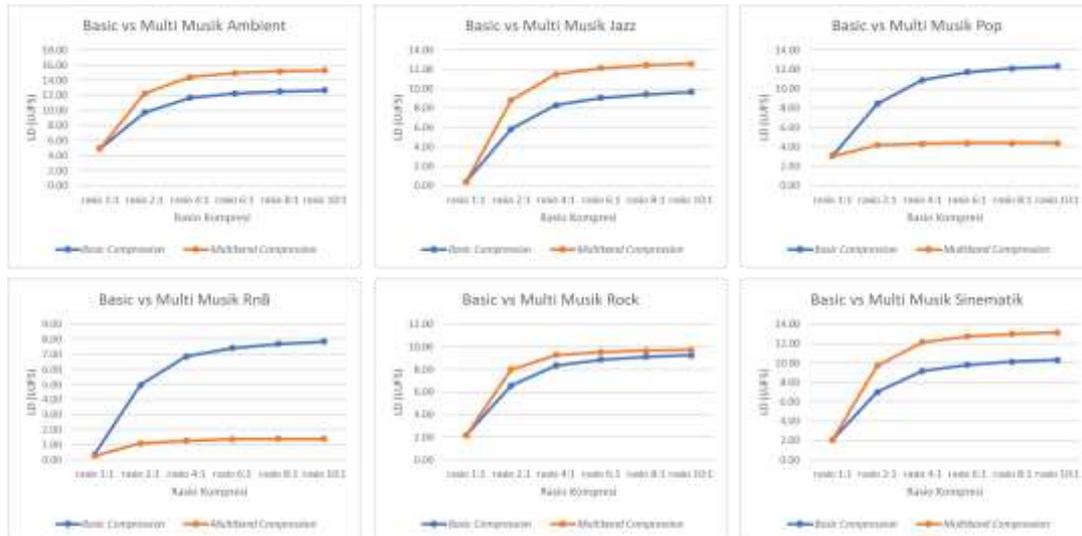
Pada gambar 4 tersebut diketahui bahwa rata-rata nilai rasio berpengaruh terhadap LD. Namun terdapat perbedaan yang signifikan antara nilai rasio 1:1 dengan rasio lainnya, nilai rasio 1:1 menurut para profesional adalah tidak terjadi adanya kompresi yang seharusnya tidak adanya LD atau perbedaan antara dialog/latar depan dengan musik/latar belakang adalah 0, hal ini terjadi karena dalam penghitungan IL (*Integrated Loudness*) menggunakan ITU-R BS.1770-4 untuk *Loudness Unit Full Scale* (LUFS) dan EBU R 128 untuk *true Peak level* (dBTP) sebagai standar untuk menentukan algoritma pengukur IL.

4.5 Pengaruh Jenis Kompresi Terhadap Hasil *Loudness Difference*

Jenis kompresi menunjukkan tingkat signifikansi $0.612 > 0,05$ yang berarti jenis kompresi menjadi satu-satunya variabel yang tidak berpengaruh terhadap hasil LD. Jenis kompresi tidak berpengaruh karena dalam teknik *sidechain compression* bergantung pada nilai *threshold*, *ratio*, *attack time*, *release time*. Karena penelitian ini menggunakan nilai rasio sebagai variabel eksperimen, maka dapat dikatakan bahwa apapun jenis kompresi akan mengakibatkan penurunan tingkat kenyaringan dikarenakan memiliki parameter didalamnya seperti *threshold*, *ratio*, *attack time*, *release time* yang menentukan proses *audio ducking* terjadi.

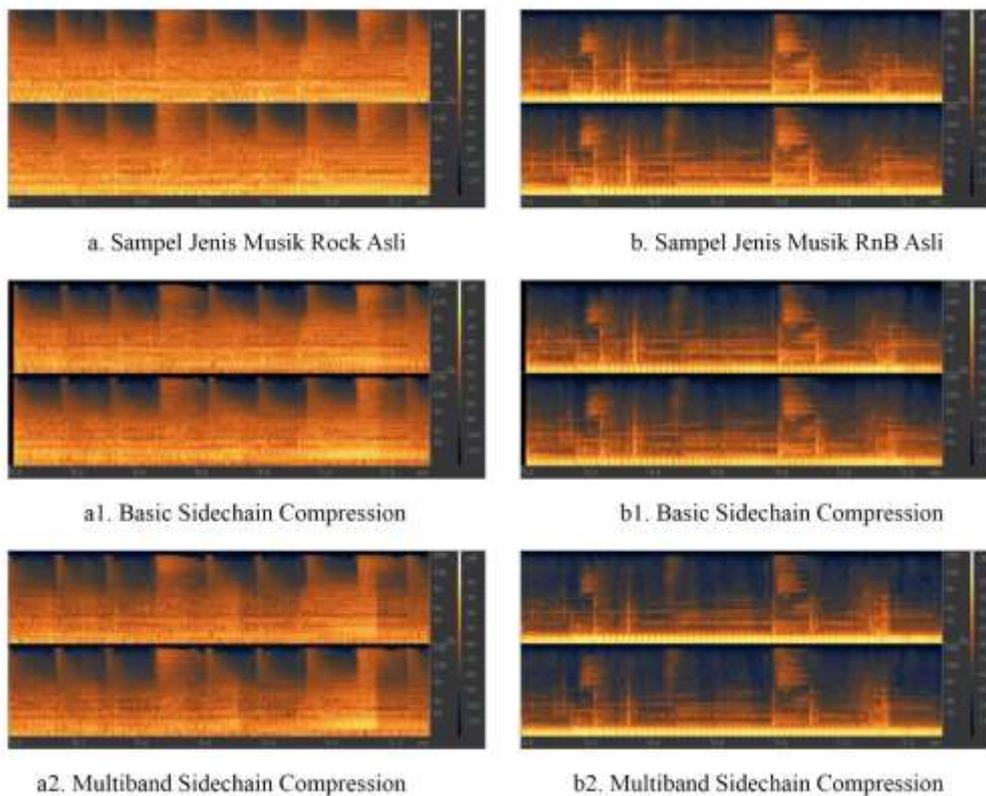
Namun jika melihat pada gambar 4 rata-rata nilai LD yang ditunjukkan kedua jenis kompresi terhadap nilai LD memiliki perbedaan bahwa nilai LD dengan jenis *Multiband Sidechain Compression* memiliki nilai lebih rendah dibandingkan *Basic Sidechain Compression*. Hal ini menjadi sebuah temuan baru bagi penulis dan melakukan sedikit penggalian lebih dalam untuk membandingkan kedua teknik kompresi tersebut dengan jenis musik dan nilai rasio kemudian di dapatkan data seperti Gambar 5.

Dari gambar 5 tersebut dapat diketahui bahwa sampel jenis musik Ambient, Jazz, Rock dan Sinematik jenis kompresi *Multiband Sidechain* memiliki nilai LD yang lebih besar sehingga menurut peneliti sebelumnya nilai LD yang besar dapat meningkatkan kejelasan dialog/narasi [3]. Sedangkan jenis kompresi *Basic Sidechain* cenderung lebih kecil.



Gambar 5 Perbandingan jenis kompresi terhadap LD pada tiap jenis musik

Untuk sampel jenis musik Pop dan RnB menunjukkan bahwa *Basic Sidechain* memiliki nilai LD yang jauh lebih tinggi daripada *Multiband Sidechain*. Hal ini terjadi mengingat konsep *Basic Sidechain* adalah menurunkan tingkat kenyaringan berdasar *Loudness Unit* sedangkan *Multiband Sidechain* berdasarkan frekuensi.



Gambar 6 Spectrogram sampel jenis musik dengan nilai rasio 4:1; (a) Rock asli, (a1) Rock dengan *Basic Sidechain compression* (a2), Rock dengan *Multiband Sidechain*, (b) RnB asli, (b1) RnB dengan *Basic Sidechain compression* (b2), RnB dengan *Multiband Sidechain*. Warna yang lebih terang menunjukkan tingkat kenyaringan yang tinggi

Dari hasil *spectrogram* pada gambar 6 tersebut dapat diketahui bahwa *basic sidechain compression* menurunkan tingkat kenyaringan secara menyeluruh pada sampel audio dan dapat

dilihat bahwa warna dalam spectrogram (a1 dan b1) lebih redup daripada sampel asli. Sedangkan *multiband sidechain compression* menurunkan tingkat kenyaringan berdasarkan frekuensi, dalam pengaturan *multiband sidechain* frekuensi yang akan dikurangi adalah 175Hz – 10kHz, dapat dilihat dalam *spectrogram* (a2 dan b2) pada frekuensi 175Hz – 10kHz lebih redup daripada sampel asli. Pada (b2) terlihat sangat jelas frekuensi dibawah 175Hz warna lebih menyala bahkan hampir sama dengan sampel asli. Faktor inilah yang menjadi alasan pada jenis musik pop dan RnB terjadi perbedaan hasil LD dengan jenis musik lainnya.

4.5 Jenis Musik dan Nilai Kompresi yang Direkomendasikan

Berdasarkan acuan dari peneliti sebelumnya yaitu Rekomendasi LD untuk *Commentary over Music* adalah 10-15 LU[3]. Dapat diketahui dengan mengeliminasi hasil LD yang dihitung seperti pada Tabel 5 sebagai berikut

Tabel 5 Rekomendasi Jenis Musik dan Nilai Rasio

Metode	Rasio Kompresi	Jenis Musik					
		Ambient	Jazz	Pop	RnB	Rock	Sinematik
Basic Sidechain Compression	Rasio 1:1	4.94	0.40	3.11	0.37	2.16	2.08
	Rasio 2:1	9.72	5.86	8.47	4.96	6.58	7.02
	Rasio 4:1	11.67	8.32	10.91	6.86	8.36	9.17
	Rasio 6:1	12.24	9.07	11.71	7.42	8.87	9.82
	Rasio 8:1	12.50	9.44	12.10	7.69	9.12	10.13
	Rasio 10:1	12.65	9.65	12.33	7.84	9.26	10.32
Multiband Dynamic Compression	Rasio 1:1	4.94	0.40	3.03	0.27	2.16	2.08
	Rasio 2:1	12.24	8.83	4.19	1.11	8.00	9.75
	Rasio 4:1	14.41	11.50	4.35	1.28	9.28	12.15
	Rasio 6:1	14.93	12.15	4.39	1.37	9.55	12.74
	Rasio 8:1	15.16	12.43	4.40	1.38	9.66	13.01
	Rasio 10:1	15.29	12.59	4.41	1.39	9.73	13.16

Pada Tabel 5 tersebut dapat diambil rekomendasi penggunaan metode kompresi terbaik yang dapat digunakan oleh penyuluh suara dalam proses *audio ducking* pada paska produksi sesuai dengan rekomendasi nilai LU pada penelitian sebelumnya berdasarkan jenis musik yang digunakan. Untuk pencampuran suara standar yang baik secara estetika dan memiliki tingkat ucapan yang jelas, direkomendasikan LD minimal 10 LU untuk *Commentary over Music* (CoM) dan 15 LU untuk *Commentary over Ambience* (CoA)[3].

Berdasarkan hasil pengukuran LD tersebut diatas penggunaan metode kompresi yang paling baik untuk (CoM) dengan asumsi LD minimal mendekati 10 LU, untuk jenis music ambient dapat menggunakan *basic side compression* dengan rasio kompresi 4:1, jenis musik jazz menggunakan *multiband dynamic compression* dengan rasio kompresi 4:1, jenis musik pop menggunakan *basic side compression* dengan rasio kompresi 4:1, untuk jenis musik sinematik menggunakan *basic side compression* dengan rasio kompresi 8:1. Sedangkan untuk jenis musik RnB dan rock berada pada nilai dibawah minimum LU rekomendasi dapat menggunakan nilai yang mendekati 10 LU yaitu *basic side compression* dengan rasio kompresi 4:1 untuk jenis music RnB dan *multiband dynamic compression* dengan rasio kompresi 6:1.

Hasil studi ini melengkapi hasil penelitian terdahulu dimana proses *manual volume automation* digunakan pada proses *ducking* dalam paska produksi berdasarkan penurunan volume audio latar belakang saja [8], sementara hasil penelitian ini menambahkan proses *ducking* berdasarkan penurunan volume audio latar belakang dengan frekuensi tertentu yang disesuaikan frekuensi dialog/narasi yang digunakan. Selain itu, pada penelitian sebelumnya pengukuran LD menggunakan metode *basic sidechain compression* [9], [11], [12]. Sedangkan penelitian ini membandingkan dua jenis metode *sidechain compression* yaitu *basic sidechain compression* dengan *multiband dynamic sidechain compression*. Dari perbandingan kedua

metode tersebut menghasilkan rekomendasi metode *sidechain compression* mana yang terbaik sesuai jenis musik yang akan digunakan seperti yang ditunjukkan pada Tabel 5 diatas.

5. Simpulan

Berdasarkan hasil analisis data mengenai penelitian ini yang bertujuan untuk mengetahui pengaruh jenis musik, jenis kompresi dan nilai rasio terhadap *Loudness Difference* (LD) sebagai acuan tingkat kejelasan suara serta rekomendasi nilai rasio berdasarkan jenis musik yang digunakan dalam proses pencampuran suara pada paska produksi multimedia maka dapat ditarik kesimpulan Jenis musik berpengaruh terhadap LD secara signifikan sehingga perlu diperhatikan pemilihan jenis musik untuk tayangan para konten creator, nilai rasio kompresi berpengaruh terhadap LD secara signifikan, nilai LD yang lebih tinggi membuat dialog/narasi menjadi terdengar lebih jelas, jenis kompresi tidak berpengaruh secara langsung terhadap hasil LD namun dapat menjadi pilihan sesuai dengan jenis musik yang digunakan dan untuk rekomendasi jenis musik, nilai rasio dan jenis kompresi dapat dilihat pada tabel 5

Sebagai pengembangan penelitian berikutnya perlu diperhatikan adanya faktor jenis musik lainnya dengan mengelompokkan *dynamic range* sampel tersebut. Perlu dilakukan pengujian kepada responden (*Listening Test*) dengan menggunakan *multiband sidechain compression* untuk membuktikan rekomendasi penelitian sebelumnya. Pengujian dilakukan dengan memperhatikan parameter *ducking* lainnya yaitu *attack time*, *release time* dan *threshold* yang bervariasi.

Daftar Referensi

- [1] Y. Kaduri, *The Oxford Handbook of Sound and Image in Western Art*. Oxford University Press, 2016.
- [2] C. L. Liu and Y. C. Chen, "Background music recommendation based on latent factors and moods," *Knowl Based Syst*, vol. 159, pp. 158–170, Nov. 2018, doi: 10.1016/j.knosys.2018.07.001.
- [3] D. Geary *et al.*, "Loudness Differences for Voice-Over-Voice Audio in TV and Streaming," *Journal of the Audio Engineering Society*, vol. 68, no. 11, pp. 810–818, Dec. 2020, doi: 10.17743/jaes.2020.0022.
- [4] R. Izhaki, *Mixing Audio*. 2011. doi: 10.1016/B978-0-240-52222-7.X0001-X.
- [5] A. P. Kefauver and D. Patschke, *Fundamentals of Digital Audio, New Edition*. A-R Editions, Inc., , 2007.
- [6] S. J. van Wijngaarden, H. J. M. Steeneken, and T. Houtgast, "Quantifying the intelligibility of speech in noise for non-native listeners," *J Acoust Soc Am*, vol. 111, no. 4, pp. 1906–1916, Apr. 2002, doi: 10.1121/1.1456928.
- [7] B. Wang and W. Wong, "Real time hearing enhancement in crowded social environments with noise gating," *Speech Commun*, vol. 99, pp. 173–182, May 2018, doi: 10.1016/j.specom.2018.03.010.
- [8] A. Torcoli, M. Torcoli, A. Freke-Morin, J. Paulus, C. Simon, and B. Shirley, "Background ducking to produce esthetically pleasing audio for TV with clear speech Title Background ducking to produce esthetically pleasing audio for TV with clear speech Background Ducking to Produce Esthetically Pleasing Audio for TV with Clear Speech," 2019. [Online]. Available: <http://usir.salford.ac.uk/id/eprint/50647/>
- [9] M. Torcoli, A. Freke-Morin, J. Paulus, C. Simon, and B. Shirley, "Preferred levels for background ducking to produce esthetically pleasing audio for TV with clear speech," *AES: Journal of the Audio Engineering Society*, vol. 67, no. 12, pp. 1003–1011, 2019, doi: 10.17743/jaes.2019.0052.
- [10] R. W. Taylor, "Towards Standardised Loudness Normalisation in Music Streaming Expanding the Genetic Landscape of Mitochondrial Diseases View project Towards Standardised Loudness Normalisation in Music Streaming," 2017, doi: 10.13140/RG.2.2.17353.03680.
- [11] Y. Tang, B. M. Fazenda, and T. J. Cox, "Automatic speech-to-background ratio selection to maintain speech intelligibility in broadcasts using an objective intelligibility metric," *Applied Sciences* , vol. 8, no. 1, p. 59, 2018, doi: 10.3390/app8010059.
- [12] M. Torcoli, J. Herre, J. Paulus, C. Uhle, H. Fuchs, and O. Hellmuth, "The Adjustment / Satisfaction Test (A/ST) for the Subjective Evaluation of Dialogue Enhancement," 2017. [Online]. Available: <http://www.aes.org/e-lib>

- [13] A. TD1004.1.15-10, *Recommendation for Loudness of Audio Streaming and Network File Playback*, 1st ed. New York: Audio Engineering Society, 2015.
- [14] S. Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktik*, Ed. Rev. VI, Cet. 14. Jakarta: Rineka Cipta, 2011.
- [15] Widodo, *Metodologi penelitian : populer & praktis* . Jakarta: Rajawali Persada , 2017.
- [16] F. N. Kerlinger, *Foundation of behavioral research*. Yogyakarta: Gadjah Mada University Press, 2006.
- [17] I. Radiocommunication Bureau, "Algorithms to measure audio programme loudness and true-peak audio level BS Series Broadcasting service (sound)," 2015. [Online]. Available: <http://www.itu.int/ITU-R/go/patents/en>
- [18] R. Zhao, S. Jing, W. Xian, and X. Wang, "Study of Speech Digital Signal Processing based on Matlab," *Proceedings of the 3rd International Conference on Computer Engineering, Information Science & Application Technology (ICCIA 2019)*, vol. 90, 2019, doi: 10.2991/iccia-19.2019.81.
- [19] I. GHOZALI, *Aplikasi Analisis Multivariate SPSS 23*. Badan Penerbit Universitas Diponegoro, 2016.